Reconciling simulated and observed views of clouds: MODIS, ISCCP, and the limits of instrument simulators

Robert Pincus *

University of Colorado and NOAA/Earth System Research Laboratory, Boulder, Colorado, USA

Steven Platnick

Earth Sciences Division, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA

Steven A. Ackerman

Cooperative Institute for Meteorological Satellite Studies, University of Wisconsin – Madison, Madison, Wiscon

Richard S. Hemler

NOAA/Geophysical Fluid Dynamics Lab, Princeton, New Jersey, USA

Robert J. Patrick Hofmann

University of Colorado and NOAA/Earth System Research Laboratory, Boulder, Colorado, USA

Corresponding Author Address: Robert Pincus, Cooperative Institute for Research in Environmental Sciences, University of Colorado/NOAA Earth Systems Research Laboratory Physical Sciences Division, 325 Broadway, R/PSD1, Boulder, CO, 80305. email: Robert.Pincus@colorado.edu.

ABSTRACT

The properties of clouds that may be observed by satellite instruments, such as optical depth and cloud top pressure, are only loosely related to the way clouds are represented in models of the atmosphere. One way to bridge this gap is through "instrument simulators," diagnostic tools that map the model representation to synthetic observations so that differences between simulator output and observations can be interpreted unambiguously as model error. But simulators may themselves be restricted by limited information available from the host model or by internal assumptions. This paper considers the extent to which instrument simulators are able to capture essential differences between MODIS and ISCCP, two similar but independent estimates of cloud properties. The authors review the measurements and algorithms underlying these two cloud climatologies, introduce a MODIS simulator, and detail data sets developed for comparison with global models using ISCCP and MODIS simulators. In nature MODIS observes less mid-level cloudiness than ISCCP, consistent with the different methods used to determine cloud top pressure; aspects of this difference are reproduced by the simulators running in a climate model. But stark differences between MODIS and ISCCP observations of total cloudiness and the distribution of cloud optical thickness can be traced to different approaches to marginal pixels, which MODIS excludes and ISCCP treats as homogeneous. These pixels, which likely contain broken clouds, cover about 15% of the planet and contain almost all of the optically thinnest clouds observed by either instrument. Instrument simulators can not reproduce these differences because the host model does not consider unresolved spatial scales and so can not produce broken pixels. Nonetheless, MODIS and ISCCP observation are consistent for all but the optically-thinnest clouds, and models can be robustly evaluated using instrument simulators by excluding ambiguous observations.

1. Evaluating simulations of present-day cloudiness with satellite observations

The fidelity of global climate models is frequently judged by comparing simulations of the present day with observations (e.g. Gleckler et al. 2007; Reichler and Kim 2008). Evaluating the distribution of clouds produced by these models (e.g. Pincus et al. 2008) is of particular interest because of the strong links between the cloud response to climate change and the overall climate sensitivity (e.g. Webb et al. 2006; Soden and Held 2006). Satellite observations are central to this task because of the near-global view of cloudiness they provide.

The evaluation of cloud properties is more complicated than that of, say, surface temperature, because the cloud properties that are observable by satellite-based instruments are quite different from the variables used to represent clouds within a model. One way to narrow this gap is to build a diagnostic tool that converts the model state into synthetic observations. The "ISCCP simulator" (Klein and Jakob 1999; Webb et al. 2001), for example, uses the internal model representation of cloudiness to produce estimates of the cloud top pressure and optical thickness that would be reported by a particular satellite observing program. The simulator accounts for effects at the pixel scale including the screening of clouds low in the atmosphere by clouds above them, the interpretation of measurements as if they arise from clouds in a single homogeneous layer, and the estimation of cloud top pressure based on infrared brightness temperatures. The treatment of pixel-scale observations is similar to, though less rigorous than, the observation operators used in data assimilation. Because climate models are compared to observations on a statistical basis, the simulator also reproduces the averaging strategies adopted during the processing of ISCCP observations.

The ISCCP simulator has proved quite valuable in diagnosing GCM performance (see, for example, Zhang et al. (2005) and Williams and Webb (2009), among many others). This has inspired a number of other simulators for cloud-related instruments, including space-borne radars (CloudSat; Haynes et al. 2007), lidars (CALIPSO; Chepfer et al. 2008), and multi-angle radiometers (MISR; see Marchand and Ackerman 2010). CALIPSO and MISR have also produced summary datasets against which the results of the simulator may be compared (Chepfer et al. 2010; Marchand et al. 2010). What these efforts share is the desire to put model predictions and observations on the same footing so that differences between the two can be unambiguously interpreted as model error.

But not all ambiguities can be removed. Here we investigate the large impact that differing treatments of uncertain

observations can have on the climatological distribution on cloud properties, demonstrate why these uncertain observations can not be modeled by instrument simulators, and show that comparisons between models and observations that exclude such observations are nonetheless robust. The next section describes long-term observations obtained from two similar but independent satellite records (ISCCP and MODIS), while Section 3 explores the most prominent differences between these two views of the Earth's clouds. Section 4 describes a method for emulating MODIS observations within a climate model while section 5 illustrates the degree to which differences between these two sets of observations are captured by instrument simulators. We conclude by discussing the roles and limitations of instrument simulators in enabling model evaluation and diagnosis.

2. Global observations of cloudiness: ISCCP and MODIS

There are at least nine available climatologies of cloud properties based on satellite observations (Stubenrauch et al. 2009). We focus on two - ISCCP and MODIS, described below - with long data records, good spatial sampling, and the ability to detect clouds throughout the atmosphere. MODIS and ISCCP estimate many of the same quantities but, as we will show, often arrive at different answers. In this section we review the way these two cloud climatologies are constructed, including descriptions of individual pixel-level retrievals and the strategies used to aggregate the observations in space and time. Our goal is to highlight the similarities and differences in these data sets and understand the consequences for model evaluation, including motivating decisions made in constructing the MODIS simulator described in Section 4. We limit ourselves to observations made during daytime as these contain a richer set of retrievals and better accuracy in cloud detection. We recommend Marchand et al. (2010) for somewhat lengthier explanations and several illuminating case studies; many of the conclusions as we reach in this section are also evident in that paper.

a. ISCCP observations

The International Satellite Cloud Climatology Project (ISCCP; see Rossow and Schiffer 1991) has produced a long (26 years and growing), well-documented, and well-used record of cloudiness. ISCCP obtains and calibrates observations at two wavelengths from operational sensors aboard geostationary and polar-orbiting satellites, then

interprets these observations to determine which pixels are cloudy and estimate the cloud top pressure p_c of the highest cloudy layer and column-integrated optical thickness τ within each (Rossow and Schiffer 1999).

ISCCP uses observations in the visible (approximately 0.6 μm, nominally 1 km resolution) and infrared (11 μm, 1-4 km resolution depending on the instrument) portions of the spectrum. Clouds are detected by comparing the observed values with estimates of the clear-sky contribution; pixels which are somewhat colder and/or brighter than the clear sky is expected to be are flagged as cloudy. Retrievals assume that each pixel is entirely covered with homogeneous clouds. Cloud top temperature is estimated from the infrared brightness temperature, from which cloud top pressure p_c is computed by comparing the temperature with a profile obtained from sounding instruments. Cloud optical thickness is inferred by assuming a thermodynamic phase based on the brightness temperature, then interpreting the observed visible intensity as an optical depth based on pre-computed tables assuming constant particle size distributions. If the cloud is optically thin the infrared emissivity is adjusted and a new cloud top temperature and pressure inferred; if no self-consistent solution can be found the cloud is assigned a pressure just above the tropopause.

Pixel-level retrievals are aggregated on an equal-area grid with roughly 250 km resolution eight times per day. The ISCCP project reports monthly averages for each of the eight-daily observations times (i.e. the monthly average diurnal cycle) and for the daily mean. (Averaging uniformly over time is equivalent to assuming that the distribution of cloud properties is well-observed at each hour of each month.) The mean optical thickness reported is not the average of the individual values, but rather the "radiative mean" τ_r defined as the value of τ that produces the mean albedo. (ISCCP provides tables that map τ to albedo and vice-versa.) It is possible to reconstruct the linear mean τ using the mean liquid and ice water paths and assumed particle sizes. ISCCP also provides a joint histogram of the cloud fraction as a function of τ and p_c; this histogram has six categories of τ and seven of p_c in the 3-hourly observations, though this is reduced to three categories in each parameter in the monthly-average files. Night-time cloud amounts are adjusted to correct for systematic differences between the day-time and night-time algorithms; these adjustments are applied to the monthly mean cloud amounts.

ISCCP processes observations from both geostationary and polar-orbiting satellites, preferring geostationary observations where they are available for more uniform temporal sampling and more stable calibration. But because cloud detection and cloud property retrievals depend on the zenith angle at which the scene is viewed (e.g. Rossow and Garder 1993; Maddux et al. 2010) the ISCCP record contains modest but significant spatially-dependent artifacts

(Norris 2000; Evan et al. 2007) visible as radial patterns centered on the sub-satellite point below the geostationary platforms. These patterns are unique to ISCCP since no other global satellite cloud climatology exploits geostationary observations.

1) ISCCP OBSERVATIONS FOR CLIMATE MODEL EVALUATION

We have constructed a special-purpose climatology of ISCCP cloud properties to facilitate comparisons with MODIS and with global models using the ISCCP simulator. We begin with the original eight-times-daily, gridded observations. We average all day-time values within a month at once. We provide two estimates of mean cloud albedo, cloud top pressure, and cloud top temperature in each grid cell: one is a linear average over time and the second is weighted by the cloud fraction at each time, consistent with the MODIS averaging strategy described in the next section. Albedo is computed using an analytic approximation to the ISCCP look-up tables. We have included the full resolution joint histogram $c_f(\tau, p_c)$ to complement the reduced-resolution joint histogram distributed by ISCCP. Monthly-mean files in NetCDF format may be obtained at `http://climserv.ipsl.polytechnique.fr/cfmip-obs.html`; we also plan to distribute them on the Earth System Grid alongside climate model results from the upcoming CMIP5 experiment (`http://cmip-pcmdi.llnl.gov/cmip5/`).

b. MODIS observations

Cloud observations from MODIS provide a useful complement to the ISCCP record. There are two MODIS (Moderate Resolution Imaging Spectrometer) instruments, one each aboard the Terra and Aqua satellites. Some of the 36 spectral observations available allow for different approaches for retrieving p_c, so MODIS retrievals contain different information and have a different set of observational biases than ISCCP. Other channels allow for the retrieval of complementary information such as thermodynamic phase and particle size; the latter is relevant for evaluating models in which particle sizes depend on aerosol properties. (Here we refer to data products by the MODIS Science Team; these may be obtained at `http://ladsweb.nascom.nasa.gov`. A separate set of retrievals is made in support of the CERES products, as described in Loeb et al. (2005) among others. Those retrievals follow somewhat different logic and have different error characteristics.)

MODIS pixel-scale retrievals, as described in detail by Platnick et al. (2003), proceed in steps:

- *Cloud detection/masking:* Clouds are initially identified by a cloud mask that uses tests incorporating a large number of spectral bands (Frey et al. 2008). Different tests are used in different domains (i.e. over oceans, vegetation, ice, etc.). The cloud mask summarizes these tests by computing a likelihood that each pixel is clear (Ackerman et al. 1998), then assigns the result to one of four categories (confident clear, probably clear, uncertain/probably cloudy, and cloudy) based on probability boundaries 0.99, 0.95, and 0.66 for clear sky.

- *Cloud top properties:* Cloud-top pressure retrievals are initially attempted using CO_2 slicing (Menzel et al. 1983), which infers the amount of CO_2 above an emitting cloud from brightness temperature measurements in several closely-spaced bands near 15 μm. (Pressure can be inferred from the amount of carbon dioxide because the gas is well-mixed.) The method fails when the integrated amount of carbon dioxide becomes optically thick and clear-sky emission dominates the signal. This limits CO_2 slicing to values of p_c less than about 700 hPa; for clouds below this level p_c is retrieved by matching the brightness temperature at 11 μm to a temperature profile (this is similar to the ISCCP algorithm). Cloud top pressure is computed at 5 km scale based on 1 km pixels; it includes all such regions in which more than 4/25 = 16% of the underlying 1 km pixels are deemed cloudy or probably cloudy by the cloud mask.

- *Cloud thermodynamic phase and optical properties:* Optical thickness τ and effective radius r_e are retrieved by minimizing the difference between the observed intensity in one visible and one near-infrared wavelength (nominally 0.86 and 2.13 μm, both at 1 km scale) and forward calculations tabulated across this parameter space. The thermodynamic phase of each pixel is required because the retrievals use separate forward calculations (i.e. lookup tables) for liquid and ice clouds. This determination is made based on a range of tests in the visible, near-infrared, and infrared portions of the spectrum; where these tests disagree the pixel is labeled "unknown/mixed phase" and liquid water cloud libraries are used in the retrieval. When the simultaneous retrieval of τ and r_e fails (i.e. the observations can not be fit to the forward calculations) the pixel is not included in aggregated statistical summaries. (For these pixels a retrieval of optical thickness alone is attempted using the fixed particle sizes assumed by ISCCP and is available in the pixel-level data.)

  Uncertainty estimates are computed for each retrieval of τ and r_e. These estimates include the effects of un-

certainties in instrument calibration and nominal plane-parallel forward models, surface spectral albedo, and spectral atmospheric correction (primarily due to above-cloud atmospheric moisture uncertainty). Other error terms can be important on a pixel-level basis (e.g., vertical and/or horizontal inhomogeneity) so the uncertainty estimates provided are a lower bound on the true uncertainty. Uncertainties are calculated from the cloud reflectance look-up tables used in the retrievals and depend on viewing and illumination geometry.

Beginning with the "Collection 5" processing algorithms introduced in 2006, MODIS retrievals identify pixels that are not likely to be completely cloud-covered (see `http://modis-atmos.gsfc.nasa.gov/C005_Changes/C005_CloudOpticalProperties_ver311.pdf`). This identification is based on tests for spectral and spatial uniformity, to identify dust, smoke, snow or ice, some partially-cloudy pixels, and sunglint, and on cloud edge detection, to remove partially-cloudy pixels. Tests suggest that the edge detection test is the most frequent trigger. In Collection 5, retrievals are not performed for pixels identified by this "clear-sky restoral" algorithm.

MODIS pixel-scale observations are aggregated on a 1° equal-angle grid (King et al. 2003). Observations from Terra and Aqua are aggregated separately. All observations from a given platform within a day are included, so the frequency of observations increases with distance from the equator, and observations from overpasses at different times (with potentially very different viewing and illumination geometries) may be included in the statistics. Statistics are weighted over time by the number of observations on each; for cloud-related variables, this means that cloudier days count more than less-cloudy days. Cloud optical properties are summarized for all clouds and separately for ice- and liquid-phase clouds.

Two estimates of cloud fraction are produced: one counts the proportion of pixels deemed by the cloud mask to be cloudy or probably cloudy, and the other counts the proportion of available pixels for which cloud optical properties are successfully retrieved. Differences between these two estimates include the proportion of pixels for which retrievals are unsuccessful but are dominated by pixels removed by the clear-sky restoration algorithm.

Averages are reported for both optical depth (i.e. $\overline{\tau}$) and its base-10 logarithm $\overline{\log_{10}(\tau)}$. Mean albedo $\overline{A}$ can, to a good approximation, be computed as $\overline{A} \approx A(10^{\overline{\log_{10}(\tau)}})$, making $\tau_l = 10^{\overline{\log_{10}(\tau)}}$ comparable to the radiative mean optical thickness τ_r reported by ISCCP.

Liquid and ice water path are computed at the pixel scale, assuming that clouds are vertically homogeneous, then

averaged separately for liquid and ice clouds.

We estimate the uncertainty in the mean of τ and r_e from pixel-level uncertainties by assuming that errors are perfectly correlated within each each grid cell on a given day but perfectly uncorrelated from day to day. This has the effect of greatly reducing typical daily uncertainties in a grid cell (by a factor of about $\sqrt{30}$ if the same number of cloudy pixels is present during a month and cloud properties are reasonably consistent over the month). True errors are almost certainly not uncorrelated from day to day — calibration and modeling errors, in particular, are probably not well represented as random errors — but neither are errors within each day likely to be perfectly correlated. Uncertainty estimates are still the subject of active development and are intended as rough guidance rather than definitive estimates.

1) MODIS OBSERVATIONS FOR CLIMATE MODEL EVALUATION

We have developed a data set of MODIS observations targeted at the evaluation of clouds in climate models. We extract a small subset (listed in Table 1) of the many data sets available from the gridded monthly files produced by the MODIS project and reformat them as NetCDF files that include the "CF" (climate and forecast; `http://cf-pcmdi.llnl.gov/`) metadata widely used in the climate community. The joint histogram of cloud top pressure and cloud optical thickness included here uses the same bin definitions as does the analogous histogram produced by ISCCP. A second joint histogram with finer resolution is available in the original data. We produce separate files for the Terra (morning) and Aqua (afternoon) platforms along with an estimate that combines observations from the two platforms, weighting by the number of pixel-level observations. New files are added as data from the operational MODIS system become available and will be updated to reflect algorithmic updates as the MODIS record is reprocessed. The data are available at `http://ladsweb.nascom.nasa.gov/` and are mirrored at `http://climserv.ipsl.polytechnique.fr/cfmip-obs.html` and on the Earth System Grid.

3. Understanding differences between ISCCP and MODIS climatologies of cloudiness and cloud properties

MODIS retrievals exploit the diversity of spectral information to offer a wider range of observations than ISCCP but many MODIS observations (see Table 1) have analogs in the ISCCP climatology. In this section we compare the climatology of cloudiness, τ, and p_c in these two data sets. Because cloud detection and the retrieval of optical thickness follow parallel approaches in the two data sets one might expect MODIS and ISCCP climatologies of cloud fraction and the distribution of optical depth to be similar. But, as we will show, differences between these two data sets can be large and arise from decisions made during pixel-scale retrievals and aggregation. Here we compare long-term means: July 1983 - June 2007 for ISCCP and July 2002 - June 2010 for MODIS.

a. Cloudiness (cloud fraction)

Cloud identification defines the population of pixels for which other cloud properties are determined so comparisons of cloud properties between datasets can be expected to agree only inasmuch as does cloud fraction. The differentiation between clear and cloud-affected pixels is consistent between MODIS and ISCCP despite the much richer range of spectral information available to MODIS: daylight cloud fractions from ISCCP are quite similar to those from the MODIS cloud mask, as illustrated in Figure 1. But the difference between MODIS cloud mask and retrieval cloud fractions is striking, exceeding 20% in some subtropical latitude bands (see also Figure 8 in Marchand et al. 2010). (The retrieval cloud fraction includes all thermodynamic phases: ice, liquid, and undetermined, as described above.) This implies that many fewer pixels are used in retrievals of cloud optical properties for MODIS than are used for ISCCP. As we discussed in Section 2b, most of the pixels without cloud property retrievals are those at the edges of cloud regions removed during clear-sky restoral. The implications of this choice will become clear in section c below.

The "Taylor diagram[1]" (Taylor 2001) in Figure 2 shows the differences between these observational data sets in

[1]Taylor diagrams indicate agreement with respect to a reference data set. The radial axis denotes the standard deviation of each data set and the azimuthal axis the space-time correlation coefficient between each test data set and the reference, and the size of each marker shows the bias with respect to the reference data set; the magnitude of the bias can be measured against the radial axis. The dashed blue line indicates the standard deviation of the reference data set; perfect agreement would be shown as a point (zero bias) where the dashed line intersects the horizontal axis

the way climate model simulations are normally evaluated against observations (see, e.g., Gleckler et al. 2007; Pincus et al. 2008): by summarizing the components of the total RMS difference between multiple data sets and a single reference. This example summarizes agreement in the composite annual cycle in cloud fraction with our ISCCP daytime-only climatology. (Here and in subsequent comparisons over time and space all data sets are mapped onto a 2.5° equal-angle grid.) Although the diurnal cycle of cloudiness can be large in some locations (e.g. Cairns 1995) temporal sampling does not strongly affect cloudiness estimates at the global scale: both the ISCCP-provided diurnal average (labeled "D2' in the figure, and including corrections for night-time cloud detection alluded to in section 2a) and the ISCCP daytime climatology sampled only at MODIS overpass times ("Resamp") are in very good agreement with the daytime-only dataset. We infer that the somewhat greater disagreement between ISCCP and the MODIS cloud mask reflects differences in performance in thin clouds, where the many spectrally-dependent tests used by MODIS may have a different outcome than the two-threshold tests used by ISCCP, and other sampling differences including the more uniform sampling of viewing and illumination angles by MODIS. Results for the MODIS retrieval fraction are omitted because the large bias dominates the figure, but the space-time standard deviation of this field and its correlation with ISCCP are about the same as the estimate from MODIS cloud mask.

Thus MODIS cloud properties are derived from two distinct populations: those detected by the cloud mask, counted in the cloud mask fraction, and included in retrievals of cloud top pressure, and the subset of these that remain after clear-sky restoral, from which the retrieval fraction and statistics of optical thickness, particle size, and liquid and ice water path are computed. Because ISCCP estimates of cloudiness are consistent with MODIS cloud detection, differences between MODIS and ISCCP estimates of cloud top pressure can be attributed to sensor capabilities and algorithmic differences, while differences in optical properties must be due to systematic relationships between those pixels included in ISCCP aggregations but excluded from MODIS.

b. Cloud top pressure

The estimation of cloud top pressure is where MODIS and ISSCP rely most heavily on different information, with ISCCP using infrared window brightness temperature to infer pressure and MODIS using CO_2 slicing. Because

(correlation one and matching standard deviations). Arcs surrounding this point denote constant root-mean-square error. Points close to one another in a Taylor diagram are equally far from the reference data set but not necessarily close to one another.

the population of clouds for which cloud top pressure is retrieved is roughly the same for MODIS and ISCCP, large differences in these estimates reflect errors in one or both of the retrievals. In practice, the largest differences between the two data sets are for mid-level clouds ($440 < p_c < 680$ hPa), which are much less frequent in the MODIS climatology than in ISCCP (see Figure 3). This is likely due to ambiguities in the ISCCP algorithms: thermal emission techniques may interpret the emission from high, cold, thin clouds over low, warm, brighter clouds as being a single layer of mid-level clouds (see, e.g., Mace et al. 2006), while CO_2 slicing is not susceptible to this error. Low clouds are correspondingly more frequent in the MODIS cloud mask than in ISCCP, indicating that than majority of pixels from the large, cloud-affected population that are described by ISCCP as mid-level are assigned higher cloud top pressures by MODIS. In the population remaining after clear-sky restoral, however, the proportion of low clouds is roughly similar in ISCCP and MODIS while the amount of high cloudiness is unaffected, implying that clear-sky restoral is more likely to exclude low clouds than high clouds. Since edge detection is the primary trigger for clear-sky restoral, we infer that low clouds are more spatially inhomogeneous than high clouds on the scales observed by MODIS (see also Maddux et al. 2010).

c. Optical thickness

Figures 1 and 2 demonstrate that the population of pixels included in ISCCP retrievals is roughly the same population of clouds identified by the MODIS cloud mask, while the population for which MODIS performs and aggregates retrievals is smaller (and, in some regions, substantially smaller). As we have said, the smaller population excludes pixels flagged by the clear-sky restoration algorithm or for which simultaneous retrievals of τ and r_e fail. The latter are rare, and off-line experiments with the MODIS processing system indicate that most pixels excluded during the clear-sky restoration are identified as being near cloud edges.

But clear sky restoral has a strong impact on the distribution of optical depths: MODIS clouds are substantially more optically thick, on average, than those observed by ISCCP (Figure 4). (These differences are far larger than the degree of approximation involved in comparing τ_l with τ_r.) This implies that the pixels identified by the MODIS clear-sky restoration algorithm tend to be classified by ISCCP as optically thin. In fact, the majority of the disagreement arises for clouds with $0.3 \leq \tau < 1.3$, which cover 1.7% in the MODIS record and 14.2% in the ISCCP record. These pixels are assigned cloud top pressures throughout the atmosphere by ISCCP (Figure 5) and categorized as low clouds

by MODIS (compare the center and right-most columns of Fig. 3). More than a third of these clouds are assigned the lowest possible cloud top pressure by ISCCP consistent with (though not conclusive proof of) ISCCP being unable to find self-consistent solutions for τ and p_c from visible and infrared observations. The optical thickness retrieved from partially-cloudy pixels is systematically smaller than from fully cloudy pixels (Chang and Coakley 2007); our result implies that the vast majority of optically thin clouds reported by ISCCP are the result of partially-cloud pixels.

When optically thin clouds are excluded global-mean cloud fraction estimates from the two platforms are quite similar (Figure 6). These non-trivial clouds are identified consistently in both data sets: space-time correlations between MODIS and ISCCP estimates of $c_f(x, y, t; \tau \geq \tau_{\text{min}})$ exceed 0.8 for $1.3 < \tau_{\text{min}} \leq 23$ (above which sample sizes are quite small). This is consistent with case studies showing that MODIS, ISCCP, and MISR estimates of the presence of optically thick clouds agree quite well (Marchand et al. 2010).

d. Interpreting marginal pixels

Figure 1 - 6 demonstrate that first-order differences between MODIS and ISCCP aggregate distributions of cloudiness and optical thickness hinge on the treatment of marginal pixels, and primarily those that contain broken clouds. We suggest that there is no entirely unambiguous way to interpret observations from these pixels, and that these differences represent limits in the ability to measure or even define what is meant by "cloud fraction."

All satellite retrievals, including those used by MODIS and ISCCP, relate intensity measurements to physical properties using simple models of the atmosphere. In particular, optical property retrievals assume that each pixel is completely covered by a single uniform cloud layer. Clear-sky restoral was introduced into MODIS processing to exclude pixels that are unlikely to fit this model because it was thought that the large values of effective radius frequently produced by MODIS might have had their root in retrievals made in broken clouds. (Excluding cloud edges did not, as it turns out, make a substantial difference to the average particle size because most retrievals at cloud edges failed to produce consistent retrievals of τ and r_e and so were ignored during aggregation.) Excluded pixels are quite frequent (about 16% as a global average); though the majority of these pixels are removed because they are near cloud edges they are also the great majority of optically thin ($\tau \leq 1.3$) pixels. To the extent that pixels near cloud edges are more likely to be broken at sub-pixel scales, this cautions against interpreting results from these pixels too literally.

Cloud detection algorithms require thresholds to separate clear and cloudy skies. Values for ISCCP were selected

in part to minimize differences between ISCCP-derived cloud amounts and independent estimates of this quantity (Rossow et al. 1993; Rossow and Schiffer 1999). In broken clouds, the desired threshold is that which balances "overestimates due to low spatial resolution offset by underestimates due to finite radiance threshold" (Rossow et al. 1993). MODIS thresholds were tuned differently (i.e. to provide a probability that the field of view contains some cloud, see Ackerman et al. 1998) but wind up producing results quite similar to ISCCP. Still, MODIS cloud detection is known to be biased in fields of small (sub-pixel scale) clouds (Zhao and Di Girolamo 2006), suggesting that perfect thresholds are not achievable.

But even unbiased cloud detection would not produce unbiased cloud retrievals. Cloud detection attempts to balance over- and underestimation of cloud occurrence; this would yield unbiased retrievals in partially-cloudy pixels only if retrievals were proportional to the product of the cloud fraction and the underlying property. This seems unlikely even in plane-parallel clouds, since strict linearity would require a) that the measured intensity is proportional to the product of cloud optical thickness and cloud fraction, and b) that the retrieval of cloud properties from radiance measurements is linear in intensity. Given the non-linear dependence of radiation fields on cloud properties and surface reflectance this is unlikely to be true in general.

Broken clouds are also those in which three-dimensional radiative transfer effects may be expected to affect the distribution of intensity observed from satellites, especially if the sub-pixel cloud fraction is small (Yang and Di Girolamo 2008; Evans et al. 2008). The underlying pixel size for both MODIS and ISCCP retrievals of optical thickness is approximately 1 km. Three-dimensional effects can be expected where the scale of the clouds is commensurate with or smaller than the "radiative smoothing scale" (Marshak et al. 1995); this depends local optical extinction but is of order 300 m – commensurate, in other words, with the scale of broken fields of cumulus (e.g. Koren et al. 2008). Some quantities, particularly optical depth, are not even well-defined in broken clouds: "There is a minimum horizontal scale at which the optical thickness is a meaningful quantity. This scale is imposed by the interaction of the radiation field with the inhomogeneous extinction field: *when 3D radiative effects are important, then optical thickness is not a meaningful parameter*" (Rossow et al. 2002, italics in the original). At near-infrared wavelengths the horizontal scale also depends significantly on microphysics (Platnick 2001).

Fundamentally, then, the interpretation of retrievals from partially-cloudy pixels is ambiguous, and MODIS and ISCCP treat this ambiguity differently. ISCCP identifies these pixels as cloudy and uses the optical properties retrieved

from them consistently, even though more than a third (5.1% of 14.2%) of the clouds assigned $0.3 \leq \tau \leq 1.3$ are assigned $p_c < 180$ hPa, consistent with ISCCP retrievals that did not produce self-consistent solutions for τ and p_c. ISCCP primarily provides information relevant to studies of the Earth's radiation budget and the project's success can be judged, in part, by noting that the top-of-atmosphere radiation budget can be closed to good accuracy using the cloud properties retrieved by the project (Zhang et al. 2004); even very numerous clouds with $\tau \leq 1.3$ have a modest effect on this budget. MODIS takes a more conservative approach, identifying these pixels as cloud-affected but excluding them from cloud retrievals. This approach accepts a truncation error in lieu of introducing unknown biases (caused by applying algorithms assuming plane-parallel, homogeneous clouds outside that range) that might be expected to differ by retrieved variable. Either choice can be rationalized, and we return to the implications of the choice for model evaluation below.

4. Synthesizing MODIS observations of clouds from model states

The cloud properties observable from space are only indirectly related to the internal representation of those clouds in global models. In general, clouds are represented by a time-evolving probability distribution of cloud condensate in each grid cell; the most frequent representation is the cloud fraction and mean liquid and ice water contents. Instrument simulators combine the model's internal description of clouds with information about the retrieval process to produce the statistics available in the observational record. Simulators can be divided into three parts: 1) a treatment of subgrid-scale variability; 2) the simulation of pixel-scale retrievals; and 3) the calculation of statistics.

The first step is required because any given profile of cloudiness, in all but the simplest cases, implies a distribution of possible retrievals. The ISSCP simulator introduced the idea of drawing Monte Carlo samples, known as "subcolums", from this distribution (see also Yu et al. 1996). The samples are constructed so that each subcolumn can be considered homogeneous while a large ensemble reproduces the input statistics (e.g. fractional cloudiness in each layer). The rules for doing so can be arbitrarily simple (Klein and Jakob 1999) or complex (Pincus et al. 2006) according to the assumptions made by the host model.

Steps 2 and 3 have direct analogies to the steps used to process observations (see Section 2). Step 2 is also conceptually similar to the forward operators used to predict observations of, for example, satellite radiance during

data assimilation. Instrument simulators in climate models tend to be less detailed than the operators used in data assimilation, however, in part because comparisons to observations are done statistically, rather than one-by-one.

We have developed a MODIS simulator in this framework. We assume that subcolumns are available. The software first operates on these to produce synthetic pixel-level retrievals, then a second procedure aggregates collections of retrievals to produce temporal averages and joint histograms. The simulator is integrated into the COSP (CFMIP Observation Simulator Package) software suite (Bodas-Salcedo et al. 2011) though it can easily be adapted to other contexts (e.g. cloud resolving models). The MODIS simulator is available at the COSP website `http://cfmip-obs-sim.googlecode.com`.

a. Simulating ISCCP observations

To provide context for the MODIS simulator it may be useful to review the ISCCP simulator that inspired it. In each subcolumn, the ISCCP simulator computes cloud optical thickness τ by summing the optical depth in each layer. Clear- and all-sky infrared zenith intensities are then computed using a simple radiative transfer model and user-provided cloud emissivity in each layer. Cloud top temperature T_c is inferred from τ, the top-of-atmosphere infrared intensity, and the surface temperature and emissivity. Two adjustments are made at this point. First, if a cloud top temperature can not be found that reproduces the top-of-atmosphere infrared intensity (i.e. if the retrieval fails) the cloud top temperature is set to 5 K less than the temperature at the tropopause (effectively placing the cloud top pressure at the tropopause itself). Secondly, the optical thickness of very thin clouds may be slightly increased to account for "IR-only" detection. Finally, the temperature profile of the troposphere is searched to find the pressure corresponding to the cloud top temperature, and cloud albedo determined by an analytic approximation to the tables used in processing ISCCP observations.

The ISCCP simulator aggregates pixel-scale retrievals to report grid-mean values of c_f (the fraction of subcolumns with $\tau > 0.3$), $\overline{A}$, τ_r, $\overline{T_c}, \overline{p_c}$, and the joint histogram $c_f(\tau, p_c)$; the latter contains a category for subcolumns with $\tau \leq 0.3$ that are present in the model but would be too thin to be observed.

b. Simulating pixel-scale MODIS observations

In keeping with the wider diversity of retrievals produced by MODIS instruments, the MODIS simulator requires a greater diversity of inputs than does the ISCCP simulator, including profiles of particle size for liquid and ice clouds $r_{e,l}$ $r_{e,i}$ and the corresponding liquid and ice optical depths at 0.67 μm within each layer of each subcolumn as a function of the model's vertical coordinate z. Users may opt to provide a single value of optical depth and the mixing ratios of cloud ice and liquid, in which case optical depth is partitioned by phase assuming that particles are in the geometric optics limit. Mixed phase clouds may be represented by providing values of particle size and optical depths for both phases within a grid cell. The value of cloud top pressure in each subcolumn as diagnosed by the ISCCP simulator is also required.

For each subcolumn we return a binary cloud mask and, for subcolumns deemed cloudy, the phase P (liquid, ice, or undetermined), cloud top pressure p_c, cloud optical thickness τ, and retrieved particle size r_e.

As with the ISCCP simulator, optical thickness is derived by integrating the sum of the extinction σ due to liquid and ice clouds through the depth of the atmosphere:

$$\tau = \int_{\text{TOA}}^{\text{sfc}} \sigma_l(z) + \sigma_i(z) dz \tag{1}$$

where the extinction is computed by dividing the optical depth in each layer by the physical thickness of the layer. This is implemented as a sum of the optical depths in each layer and, like the ISCCP simulator, assumes that optical thickness can be retrieved with no error.

We mimic the daytime cloud mask be specifying a minimum detectable optical thickness τ_{min}. In the present version we set $\tau_{\text{min}} = 0.3$ based on comparisons between MODIS and high-sensitivity lidars (Ackerman et al. 2008). Note that the same threshold is used in the ISCCP simulator.

Comparisons of CO_2 slicing retrievals of p_c with profiles of extinction from space-borne lidar suggest that the pressure reported is roughly that at which the vertically-integrated extinction becomes 1 (Holz et al. 2006). We approximate the retrieval of cloud top pressure by reporting the mean extinction-weighted pressure of the first optical depth., i.e.

$$p_c = \frac{1}{\tau} \int_{\text{TOA}}^{\tau} p(z)(\sigma_l(z) + \sigma_i(z)) dz, \quad \tau = \min(1, \int_{\text{TOA}}^{0} \sigma_l(z) + \sigma_i(z) dz) \tag{2}$$

The retrieved value of cloud top pressure is not constrained to fall at a model interface level. We mimic the upper limit

of CO_2 slicing (see section 2b) by reporting the cloud top pressure from the ISCCP simulator for subcolumns where Eq. 2 provides a value greater than 700 hPa.

Similarly, calculations performed by Wind et al. (2010) suggest that MODIS phase determination is sensitive to the phase in the highest portion of the cloud. We model this simply by computing

$$P = \frac{1}{\tau}\int_{\text{TOA}}^{\tau} P_l(z)\sigma_l(z) + P_i(z)\sigma_i(z)dz, \quad \tau = \min(1, \int_{\text{TOA}}^{0} \sigma_l(z) + \sigma_i(z)dz) \tag{3}$$

where P is an integer indicating liquid (P_l) or ice (P_i). Where less than 70% of the extinction in the first optical depth arises from a single phase (i.e. where $0.3 < |P - P_l| < 0.7$) the phase is considered undetermined; otherwise P is rounded to identify the phase as entirely liquid or ice.

Particle size is estimated using a simplified pseudo-inversion. During the development of the operational MODIS algorithms we built two large lookup tables, one each for liquid and ice clouds, summarizing the optical properties of cloud particle size distributions as a function of effective radius. For use in the MODIS simulator we approximate the lookup tables for wavelength $\lambda = 2.1\mu$m with polynomial fits for the dependence of single scattering albedo ω_0 and asymmetry parameter g on effective radius r_e. We use these fits to determine the total (liquid plus ice) cloud optical properties $\omega_0(r_e(z))$ and $g(r_e(z))$, which we combine with the total extinction $\sigma_l(z) + \sigma_i(z)$ to compute a predicted top-of-atmosphere near-infrared albedo R_p for each subcolumn using the two-stream approximation. We compare this albedo to a set of albedos $R_t(r_e)$ made for homogenous clouds of phase P (as determined in Eq. 3) using the value of optical thickness obtained in Eq. 1. We compute albedo at a fixed set of trial sizes and linearly interpolate to find the value of r_e that minimizes $R_p - R_t(r_e)$. As in the processing of MODIS observations, subcolumns for which phase is undetermined assume liquid drops in the retrieval of particle size. If the pseudo-retrieval fails the values of all cloud properties are set to missing values. Uncertainties due to the clear atmosphere and surface are neglected.

c. Simulating grid-scale MODIS observations

Synthetic MODIS pixels are aggregated to produce the statistics available in the custom observational data set (see Table 1) although only a single set of cloud fractions is produced. In practice, these statistics are normally computed for each model column at each diagnostic time step. This slightly complicates time averaging: MODIS observations represent the average over the number of pixels observed, so a time series of aggregate results from the

MODIS simulator must be weighted by the appropriate cloud fraction at each time in order to produce a time-mean comparable to the observations. (The ISCCP simulator shares this trait.)

5. Understanding differences in ISCCP-like and MODIS-like views of a climate model's cloudiness

To what extent can the differences between the MODIS and ISCCP views of the earth's cloudiness described in Section 3 be captured by instrument simulators like those described in Section 4? We address this question using results from AM3 (Donner et al. 2011), a new global model developed by the NOAA Geophysical Dynamics Laboratory, in which we have implemented the full COSP package. The description of subgrid-scale structure in AM3 is fairly involved, and includes a method for diagnosing internal variability based on cloud condensate amounts and cloud fraction in each layer as well as an overlap assumption that accounts for correlations in condensate concentration (Pincus et al. 2005, 2006). The model also predicts aerosol concentration and cloud drop sizes; MODIS is the one of two global observational data sets available for evaluation of the latter. (The other is PATMOS-X; see `http://cimss.ssec.wisc.edu/patmosx/`.) Results here are taken from a 26-year run (1980 - 2006) using observed sea-surface temperatures and greatly simplified sulphate chemistry.

Comparing the descriptions of the ISCCP and MODIS simulators (Sections 4a and b, respectively), one can see that the determination of optical thickness (and hence cloud detection) in individual subcolumns is nearly the same in the MODIS and ISCCP simulators. There are two (presumably rare) exceptions: subcolumns containing very optically thin clouds, for which the ISCCP may increase the optical thickness to flag the subcolumn as cloudy, and subcolumns which are excluded because the MODIS simulator's particle size pseudo-retrieval fails. Indeed, total cloud cover produced by the ISCCP simulator running in AM3 (56.2%, counting only "detectable" clouds with $\tau > 0.3$) is only slightly larger than that produced by the MODIS simulator (54%). Almost half this difference is due to clouds with $1.3 \leq \tau < 3.6$ (see Figure 7) although simulated ISCCP cloud fractions exceed simulated MODIS values slightly across the range of optical thickness, which we attribute to failures of the MODIS simulator particle size pseudo-retrieval in mixed-phase clouds for which liquid lookup tables were used.

The degree to which the two simulators produce different distributions of cloud top pressure, on the other hand,

depends more strongly on the joint distribution of τ and p_c produced by the host model. If, for example, all high clouds in the model are optically thick then infrared-based estimates should agree well with the proxy for CO_2 slicing (Eq. 2). This is not the case with AM3, however: the largest differences between synthetic MODIS and ISCCP observations is related to high clouds (Figure 8), which the ISCCP simulator puts in the highest bin ($p_c \leq 180$ hPa) and the MODIS simulator places somewhat lower in the atmosphere. This result is somewhat surprising at face value since both simulators assume perfect knowledge of τ, but may arise either from tenuous high clouds or from the failure of the ISCCP simulator to identify cloud top pressure from cloud top temperature for many optically thin clouds: comparing the joint histogram of cloud top pressure and optical thickness (Figure 9) for the two simulators are quite similar for clouds with $\tau > 3.6$, while the ISCCP simulator puts essentially all clouds with $\tau \leq 1.3$ in the highest cloud top pressure category.

Figures 7 and 8 also show observed distributions of cloudiness and Fig. 9 can be compared with Fig. 5. Robust differences between the model and observations, such as the lack of modeled clouds with $p_c > 800$ mb, can be confidently attributed to model error. For clouds with $\tau \leq 1.3$, however, the difference between the observations is commensurate with the difference between the model and either set of observations.

6. Interpreting differences among observations and model estimates of cloud properties

Model evaluation relies on the comparison of simulations with observations, and where some aspect of a model's state is not directly observable a further model of the observation process must be invoked. Instrument simulators are one class of such models and the evaluation of clouds in global models would be substantially less direct in their absence. The term "instrument simulator" is a misnomer since all three steps described in Section 4 are relevant: assumptions made about the vertical coherence of clouds can have a profound impact on model predictions of column-integrated quantities such as optical thickness (Morcrette and Jakob 2000; Pincus et al. 2006), while a naïve comparison of mean optical thickness between ISCCP (which reports τ_r) and MODIS (which reports $\overline{\tau}$ among other quantities) would be misleading. Some mapping between model state and observations is always necessary; instrument simulators can be thought of as the most direct mapping for satellite observations of clouds.

But, like all models, instrument simulators are necessarily incomplete. Our results suggest at least four circumstances in which even perfect instrument simulators may not reproduce artifacts in the observations.

- Some observational errors depend on the state of the clouds and atmosphere: the determination of cloud top temperature from an observation of infrared intensity, for example, is less subject to error when the cloud is optically thick. If the distribution of cloud parameters in nature differs from those in the model being evaluated even ideal simulators will reproduce a different mixture of observations and observational artifacts than is observed.

- Simulators may assume no error in some part of the retrieval process, as both the MODIS and ISCCP simulators do for the retrieval of cloud optical thickness, when errors may be present in the observations. Making this assumption, even if there are no reasonable alternatives, may have further implications. In particular, we expect that assuming perfect knowledge of τ, even for small τ, means that the ISCCP simulator is more likely to find consistent τ, T_c pairs than are the corresponding observations.

Given that the ISCCP simulator predicts cloud top pressure quite accurately given accurate inputs (Mace et al. 2011), we suspect the large discrepancies between the observed and simulated distributions of cloudiness with height (Fig. 8) is due in part to some combination of these two factors.

- Observational artifacts may be caused by factors that are not treated by the simulators. The modest but clear geographic patterns in ISCCP observations (Section 2a), for example, are due to the dependence of retrieved cloud properties on view angle and ISCCP's preference for geostationary observations. Including this effect in global models is possible but has not yet been considered important enough to warrant the considerable effort.

- Observational artifacts may be caused by factors that are conceptually incompatible with some aspect of the simulators. As we showed in Section 3, differences in the way ISCCP and MODIS treat marginal pixels have a dramatic effect on the climatology of cloudiness and the distribution of optical thickness in thin clouds. To the extent that marginal pixels are those containing broken clouds, the frequency of these pixels depends directly on the relative spatial scales of the satellite instruments and the clouds themselves. But the Monte Carlo generation of subcolumns (Sec. 4) is agnostic with respect to scale and can not produce partially-cloudy pixels. One might, for example, develop a model of cloud spatial scale that depends on cloud type (e.g. Alexandrov et al. 2010)

but the synthetic observational artifacts introduced would depend entirely on this model and so might be better addressed independently.

But though instrument simulators cannot entirely close the gap between models and observations of clouds, there are broad regimes in which they work well: we showed in Section 3 that MODIS and ISCCP agree well in all but marginal pixels and, in Section 5, that the corresponding instrument simulators share this behavior (see also Mace et al. 2011). This suggests that model evaluation can be made more robust by focusing on that part of parameter space where instrument simulators capture the bulk of the observational artifacts.

One specific implication is that comparisons of total cloud fraction between models and observations may be unenlightening. Cloud fraction has historically been attractive because it has been considered easier to make a binary decision about the presence or absence of cloud than to infer the value of some continuous quantity (see, e.g. Pincus et al. 2011). But, as the results of Section 3 highlight, total cloud fraction is quite sensitive to the spatial scale at which it is measured (see also Wielicki and Parker 1992; Di Girolamo and Davies 1997), and at the kilometer scales of current satellite instruments different assumptions can have dramatic impacts on the climatology. The agreement between MODIS and ISCCP is greatly improved when clouds with $\tau \leq 1.3$ are excluded, suggesting that applying a similar (entirely empirical) filter to the models before evaluation would make the comparisons more robust. Figure 10 demonstrates how this impacts the evaluation of AM3. The upper panel shows an evaluation of cloud fraction (ISCCP and MODIS retrieval fraction) against the respective simulators; the biases between the two observational data sets (pink) is much larger than either model estimate and a substantial fraction of the variability in cloudiness. Restricting the comparison to pixels and subcolumns with $\tau > 1.3$ brings agreement between the two sets observations to nearly the levels seen in comparing cloud-affected area (compare with Figure 2). Using more restrictive thresholds (as in Marchand et al. 2010, , for example) reduces biases differences between the MODIS and ISCCP data sets but other measures, including the bias relative to the remaining cloud fraction and the correlations between the two climatologies, are maximized by including the largest possible sample.

The stark differences between MODIS and ISCCP estimates of cloudiness suggest more flexible strategies for aggregating observations. The treatments of marginal pixels in present versions of the data lie at two ends of a spectrum. The next revision of the MODIS data will include separate aggregation of one or more categories of pixels flagged by the clear-sky restoral algorithm, allowing users to compute statistics for larger populations at the cost of

increasing uncertainty. ISCCP could follow suit by separately aggregating those pixels for which self-consistent pairs of τ and p_c could not be found. Even a small number of categories would allow users to explore the consequences of making less drastic decisions and to more sharply delineate more substantive disagreements.

Acknowledgments.

We appreciate advice from Stephen A. Klein on the design of the MODIS simulator and in helping us sharpen the arguments in this work. Gang Ye and Gary Fu of the MODIS Data Processing System went to great lengths to produce our custom dataset from the original MODIS files. We are grateful to Yuying Zhang for help in the production of the ISCCP data set and to Paul Hubanks for sharing essential knowledge about the details of the MODIS Level-3 products. RP thanks the Max-Planck-Institute for Meteorology (Hamburg) for gracious hospitality and lively discussions during portions of this project. This work was supported by NASA under grant NNX08AD65G.

REFERENCES

Ackerman, S. A., R. E. Holz, R. Frey, E. W. Eloranta, B. C. Maddux, and M. McGill, 2008: Cloud detection with MODIS. Part II: Validation. *Journal of Atmospheric and Oceanic Technology*, **25 (7)**, 1073–1086, doi:10.1175/2007JTECHA1053.1.

Ackerman, S. A., K. I. Strabala, W. P. Menzel, R. A. Frey, C. Moeller, and L. E. Gumley, 1998: Discriminating clear sky from clouds with MODIS. *Journal of Geophysical Research - Atmospheres*, **103 (D24)**, 32 141–32 157, doi:10.1029/1998JD200032.

Alexandrov, M. D., A. Marshak, and A. S. Ackerman, 2010: Cellular statistical models of broken cloud fields. Part I: Theory. *Journal of the Atmospheric Sciences*, **67 (7)**, 2125–2151, doi:10.1175/2010JAS3364.1.

Bodas-Salcedo, A., et al., 2011: COSP: satellite simulation software for model assessment. *Bulletin Of The American Meteorological Society*, to appear, doi:10.1175/2011BAMS2856.1.

Cairns, B., 1995: Diurnal variations of cloud from ISCCP data. *Atmospheric Research*, **37 (1-3)**, 133 – 146, doi:10.1016/0169-8095(94)00074-N.

Chang, F.-L. and J. A. Coakley, 2007: Relationships between marine stratus cloud optical depth and temperature: Inferences from AVHRR observations. *Journal of Climate*, **20 (10)**, 2022–2036, doi:10.1175/JCLI4115.1.

Chepfer, H., S. Bony, D. Winker, G. Cesana, J. L. Dufresne, P. Minnis, C. J. Stubenrauch, and S. Zeng, 2010: The GCM-oriented CALIPSO cloud product (CALIPSO-GOCCP). *Journal of Geophysical Research - Atmospheres*, **115**, doi:10.1029/2009JD012251.

Chepfer, H., S. Bony, D. Winker, M. Chiriaco, J. L. Dufresne, and G. Seze, 2008: Use of CALIPSO lidar observations to evaluate the cloudiness simulated by a climate model. *Geophysical Research Letters*, **35 (15)**, doi:10.1029/2008GL034207.

Di Girolamo, L. and R. Davies, 1997: Cloud fraction errors caused by finite resolution measurements. *Journal of Geophysical Research - Atmospheres*, **102 (D2)**, 1739–1756, doi:10.1029/96JD02663.

Donner, L. J., et al., 2011: The dynamical core, physical parameterizations, and basic simulation characteristics of the atmospheric component of the GFDL global coupled model CM3. *Journal of Climate*, to appear, doi:10.1175/2011JCLI3955.1.

Evan, A. T., A. K. Heidinger, and D. J. Vimont, 2007: Arguments against a physical long-term trend in global ISCCP cloud amounts. *Geophysical Research Letters*, **34**, L04 701, doi:10.1029/2006GL028083.

Evans, K. F., A. Marshak, and T. Várnai, 2008: The potential for improved boundary layer cloud optical depth retrievals from the multiple directions of MISR. *Journal of the Atmospheric Sciences*, **65 (10)**, 3179–3196, doi:10.1175/2008JAS2627.1.

Frey, R. A., S. A. Ackerman, Y. Liu, K. I. Strabala, H. Zhang, J. R. Key, and X. Wang, 2008: Cloud detection with MODIS. Part I: Improvements in the MODIS cloud mask for collection 5. *Journal of Atmospheric and Oceanic Technology*, **25 (7)**, 1057–1072, doi:10.1175/2008JTECHA1052.1.

Gleckler, P. J., K. E. Taylor, and C. Doutriaux, 2007: Performance metrics for climate models. *Journal of Geophysical Research - Atmospheres*, **113**, D06 104, doi:10.1029/2007JD008972.

Haynes, J. M., R. Marchand, Z. Luo, A. Bodas-Salcedo, and G. L. Stephens, 2007: A multipurpose radar simulation package: Quickbeam. *Bulletin Of The American Meteorological Society*, **88 (11)**, 1723+, doi:10.1175/BAMS-88-11-1723.

Holz, R. E., S. A. Ackerman, P. Antonelli, F. Nagle, R. O. Knuteson, M. McGill, D. L. Hlavka, and W. D. Hart, 2006: An improvement to the high-spectral-resolution CO_2-slicing cloud-top altitude retrieval. *Journal of Atmospheric and Oceanic Technology*, **23 (5)**, 653–670, doi:10.1175/JTECH1877.1.

King, M. D., et al., 2003: Cloud and aerosol properties, precipitable water, and profiles of temperature and water vapor from MODIS. *IEEE Transactions on Geoscience and Remote Sensing*, **41 (2)**, 442–458, doi:10.1109/TGRS.2002.808226.

Klein, S. A. and C. Jakob, 1999: Validation and sensitivities of frontal clouds simulated by the ECMWF model. *Monthly Weather Review*, **127**, 2514–2531, doi:10.1175/1520-0493(1999)127⟨2514:VASOFC⟩2.0.CO;2.

Koren, I., L. Oreopoulos, G. Feingold, L. A. Remer, and O. Altaratz, 2008: How small is a small cloud? *Atmospheric Chemistry and Physics*, **8 (14)**, 3855–3864, doi:10.5194/acp-8-3855-2008.

Loeb, N. G., S. Kato, K. Loukachine, and N. Manalo-Smith, 2005: Angular distribution models for top-of-atmosphere radiative flux estimation from the Clouds and the Earth's Radiant Energy System instrument on the Terra satellite. Part I: Methodology. *Journal of Atmospheric and Oceanic Technology*, **22 (4)**, 338–351, doi:10.1175/JTECH1712.1.

Mace, G. G., S. Houser, S. Benson, S. A. Klein, and Q. Min, 2011: Critical evaluation of the isccp simulator using ground-based remote sensing data. *Journal of Climate*, to appear, doi:10.1175/2010JCLI3517.1.

Mace, G. G., et al., 2006: Cloud radiative forcing at the Atmospheric Radiation Measurement program climate research facility: 1. Technique, validation, and comparison to satellite-derived diagnostic quantities. *Journal of Geophysical Research - Atmospheres*, **111 (D11)**, D11S90, doi:10.1029/2005JD005921.

Maddux, B. C., S. A. Ackerman, and S. Platnick, 2010: Viewing geometry dependencies in MODIS cloud products. *Journal of Atmospheric and Oceanic Technology*, **27 (9)**, 1519–1528, doi:10.1175/2010JTECHA1432.1.

Marchand, R. and T. P. Ackerman, 2010: An analysis of cloud cover in multiscale modeling framework global climate model simulations using 4 and 1 km horizontal grids. *Journal of Geophysical Research - Atmospheres*, **115**, D16 207, doi:10.1029/2009JD013423.

Marchand, R., T. P. Ackerman, M. Smyth, and W. B. Rossow, 2010: A review of cloud top height and optical depth histograms from MISR, ISCCP and MODIS. *Journal of Geophysical Research - Atmospheres*, **115**, D16 206, doi:10.1029/2009JD013422.

Marshak, A., A. D. Davis, W. J. Wiscombe, and R. F. Cahalan, 1995: Radiative smoothing in fractal clouds. *Journal of Geophysical Research - Atmospheres*, **100 (D12)**, 26 247–26 261.

Menzel, W. P., W. L. Smith, and T. R. Stewart, 1983: Improved cloud motion wind vector and altitude assignment using VAS. *Journal of Climate and Applied Meteorology*, **22 (3)**, 377–384, doi:10.1175/1520-0450(1983)022⟨0377:ICMWVA⟩2.0.CO;2.

Morcrette, J.-J. and C. Jakob, 2000: The response of the ECMWF model to changes in the cloud overlap assumption. *Monthly Weather Review*, **128 (6)**, 1707–1732, doi:10.1175/1520-0493(2000)128⟨1707:TROTEM⟩2.0.CO;2.

Norris, J. R., 2000: What can cloud observations tell us about climate variability? *Space Science Reviews*, **94 (1-2)**, 375–380, doi:10.1023/A:1026704314326.

Pincus, R., C. P. Batstone, R. J. P. Hofmann, K. E. Taylor, and P. J. Gleckler, 2008: Evaluating the present-day simulation of clouds, precipitation, and radiation in climate models. *Journal of Geophysical Research - Atmospheres*, **113**, D14 209, doi:10.1029/2007JD009334.

Pincus, R., C. Hannay, S. A. Klein, K.-M. Xu, and R. S. Hemler, 2005: Overlap assumptions for assumed probability distribution function cloud schemes in large-scale models. *Journal of Geophysical Research - Atmospheres*, **110**, D15S09, doi:10.1029/2004JD005100.

Pincus, R., R. S. Hemler, and S. A. Klein, 2006: Using stochastically generated subcolumns to represent cloud structure in a large-scale model. *Monthly Weather Review*, **134 (12)**, 3644–3656, doi:10.1175/MWR3257.1.

Pincus, R., R. J. P. Hofmann, J. L. Anderson, K. Raeder, N. Collins, and J. S. Whitaker, 2011: Can fully accounting for clouds in data assimilation improve short-term forecasts? *Monthly Weather Review*, **139**, 946–957, doi:10.1175/2010MWR3412.1.

Platnick, S., 2001: Approximations for horizontal photon transport in cloud remote sensing problems. *Journal of Quantitative Spectroscopy and Radiative Transfer*, **68**, 75–99, doi:10.1016/S0022-4073(00)00016-9.

Platnick, S., M. D. King, S. A. Ackerman, W. P. Menzel, B. A. Baum, J. C. Riédi, and R. A. Frey, 2003: The MODIS cloud products: Algorithms and examples from Terra. *IEEE Transactions on Geoscience and Remote Sensing*, **41 (2)**, 459–473, doi:10.1109/TGRS.2002.808301.

Reichler, T. and J. Kim, 2008: How well do coupled models simulate today's climate? *Bulletin Of The American Meteorological Society*, **89 (3)**, 303–311, doi:10.1175/BAMS-89-3-303.

Rossow, W. B., C. Delo, and B. Cairns, 2002: Implications of the observed mesoscale variations of clouds for the earth's radiation budget. *Journal of Climate*, **15 (6)**, 557–585, doi:10.1175/1520-0442(2002)015⟨0557:IOTOMV⟩2.0.CO;2.

Rossow, W. B. and L. C. Garder, 1993: Validation of ISCCP cloud detections. *Journal of Climate*, **6 (12)**, 2370–2393, doi:10.1175/1520-0442(1993)006⟨2370:VOICD⟩2.0.CO;2.

Rossow, W. B. and R. A. Schiffer, 1991: ISCCP cloud data products. *Bulletin Of The American Meteorological Society*, **72 (1)**, 2–20, doi:10.1175/1520-0477(1991)072⟨0002:ICDP⟩2.0.CO;2.

Rossow, W. B. and R. A. Schiffer, 1999: Advances in understanding clouds from ISCCP. *Bulletin Of The American Meteorological Society*, **80**, 2261–2287, doi:10.1175/1520-0477(1999)080⟨2261:AIUCFI⟩2.0.CO;2.

Rossow, W. B., A. W. Walker, and L. C. Garder, 1993: Comparison of ISCCP and other cloud amounts. *Journal of Climate*, **6 (12)**, 2394–2418, doi:10.1175/1520-0442(1993)006⟨2394:COIAOC⟩2.0.CO;2.

Soden, B. J. and I. M. Held, 2006: An assessment of climate feedbacks in coupled ocean–atmosphere models. *Journal of Climate*, **19 (14)**, 3354–3360, doi:10.1175/JCLI3799.1.

Stubenrauch, C., S. Kinne, and the GEWEX Cloud Assessment Team, 2009: Assessment of global cloud climatologies. *GEWEX Newsletter*, **19 (1)**, 6–7.

Taylor, K. E., 2001: Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research - Atmospheres*, **106 (D7)**, 7183–7192, doi:10.1029/2000JD900719.

Webb, M. J., C. Senior, S. Bony, and J.-J. Morcrette, 2001: Combining ERBE and ISCCP data to assess clouds in the Hadley Centre, ECMWF and LMD atmospheric climate models. *Climate Dynamics*, **17 (12)**, 905–922, doi:10.1007/s003820100157.

Webb, M. J., et al., 2006: On the contribution of local feedback mechanisms to the range of climate sensitivity in two GCM ensembles. *Climate Dynamics*, **27 (1)**, 17–38, doi:10.1007/s00382-006-0111-2.

Wielicki, B. A. and L. Parker, 1992: On the determination of cloud cover from satellite sensors: The effects of sensor spatial resolution. *Journal of Geophysical Research - Atmospheres*, **D12**, 12 799–12 823, doi:10.1029/92JD01061.

Williams, K. D. and M. J. Webb, 2009: A quantitative performance assessment of cloud regimes in climate models. *Climate Dynamics*, **33 (1)**, 141–157, doi:10.1007/s00382-008-0443-1.

Wind, G., S. Platnick, M. D. King, P. A. Hubanks, M. J. Pavolonis, A. K. Heidinger, P. Yang, and B. A. Baum, 2010: Multilayer cloud detection with the MODIS near-infrared water vapor absorption band. *Journal of Applied Meteorology and Climatology*, **49 (11)**, 2315–2333, doi:10.1175/2010JAMC2364.1.

Yang, Y. and L. Di Girolamo, 2008: Impacts of 3-D radiative effects on satellite cloud detection and their consequences on cloud fraction and aerosol optical depth retrievals. *Journal Of Geophysical Research-Atmospheres*, **113 (D4)**, D04 213, doi:10.1029/2007JD009095.

Yu, W., M. Doutriaux, G. Sèze, H. Treut, and M. Desbois, 1996: A methodology study of the validation of clouds in GCMs using ISCCP satellite observations. *Climate Dynamics*, **12**, 389–401, doi:10.1007/BF00211685.

Zhang, M. H., et al., 2005: Comparing clouds and their seasonal variations in 10 atmospheric general circulation models with satellite measurements. *Journal of Geophysical Research - Atmospheres*, **110 (D15)**, D15S02, doi: 10.1029/2004JD005021.

Zhang, Y., W. B. Rossow, A. A. Lacis, V. Oinas, and M. I. Mishchenko, 2004: Calculation of radiative fluxes from the surface to top of atmosphere based on ISCCP and other global data sets: Refinements of the radiative transfer model and the input data. *Journal of Geophysical Research - Atmospheres*, **109**, D19 105, doi:10.1029/2003JD004457.

Zhao, G. and L. Di Girolamo, 2006: Cloud fraction errors for trade wind cumuli from EOS-Terra instruments. *Geophysical Research Letters*, **33 (20)**, L20 802, doi:10.1029/2006GL027088.

List of Tables

TABLE 1. Parameters available in the MODIS data set intended for climate model evaluation. All quantities are available as monthly averages on a 1° equal-angle grid. The difference between the total and (liquid + ice) retrieval cloud fractions is the fraction of detections for which phase is undetermined. Bin boundaries for the joint histogram of cloud top pressure and optical thickness follow the ISCCP definition.

Parameter	Subsets	Uncertainty provided
Cloud fraction - mask	total;	
	high, mid, low	
Cloud fraction - retrieval	total;	
	high, mid, low;	
	liquid, ice	
Mean cloud top pressure		
Mean optical thickness	total;	
	liquid, ice	y
Mean $\log_{10}$(optical thickness)	total;	
	liquid, ice	y
Mean particle size	liquid, ice	y
Mean water path	liquid, ice	y
Optical thickness/cloud top pressure joint histogram		

List of Figures

5 Climatological distribution of global-mean cloud fraction $c_f(\tau, p_c)$ as a function of cloud top pressure p_c (vertical axis) and cloud optical thickness τ (horizontal axis) observed by ISCCP (upper panel) and MODIS (lower). ISCCP provides retrievals for all pixels identified as cloud-affected, while MODIS excludes those identified as marginal by the "clear-sky restoral" process described in the text. The broken clouds most frequently removed during clear-sky restoral represent almost all the optical thinnest clouds ($\tau < 1.3$) observed by ISCCP. The frequent observation of very high, very thin clouds by ISCCP may indicate that many of these observations are difficult to interpret self-consistently. 37

6 Global-mean climatological values of day time cloud fraction $c_f(\tau \geq \tau_{\text{min}})$ as a function of the minimum optical depth τ_{min}. The global-mean value is computed by summing the joint histogram $c_f(x, y, t; \tau \geq \tau_{\text{min}}, p_c)$ and averaging in time and space. The large difference in cloudiness between MODIS retrievals, on the one hand, and ISCCP and the MODIS cloud mask, on the other (see Fig. 1) is due almost entirely to clear-sky restoral removing pixels with $0.3 \leq \tau \leq 1.3$; most of these are initially identified as low clouds by MODIS. 38

7 Marginal histogram of global-mean cloud fraction as a function of optical thickness τ in observations (solid lines) and from instrument simulators running in the AM3 climate model (dashed). ISCCP and the ISCCP simulator are in green; MODIS and the MODIS simulator are in blue. The instrument simulators produce very similar distributions, though some MODIS simulator cloud fractions are slightly smaller at most values of τ due to failed retrievals of particle size in mixed-phase clouds. Neither simulator accounts for retrievals in scattered clouds where the observations disagree most sharply (see Section 3). Clouds with $\tau \leq 0.3$ are quite common in AM3 though they are not detectable in passive observations. 39

8 Marginal histogram of global-mean cloud fraction as a function of cloud top pressure p_c in observations (solid lines) and from instrument simulators running in the AM3 climate model (dashed). ISCCP and the ISCCP simulator (excluding clouds with $\tau \leq 0.3$) are in green; MODIS and the MODIS simulator are in blue. Total observed cloud fractions are lower in MODIS than ISCCP primarily because MODIS does not aggregate marginal pixels which tend to be optically thin and low in the atmosphere. ISCCP thermal algorithms assigns far more cloudiness to the middle of the atmosphere ($p_c \leq 440 < 680$ hPa) than does MODIS's CO_2 slicing method. The MODIS and ISCCP simulators do not capture this detail, in part because the joint distribution of clouds $c_f(\tau, p_c)$ in AM3 is not the same as the distribution in nature. 40

9 Climatological distribution of global-mean cloud fraction $c_f(\tau, p_c)$ as a function of cloud top pressure p_c (vertical axis) and cloud optical thickness τ (horizontal axis) from the ISCCP (upper row) and MODIS (lower row) simulators running in the AM3 climate model. Very optically thin clouds ($\tau \leq 0.3$) cover about 11% of the earth in AM3; these are not detectable in either set of observations. The simulators reproduce the higher cloud top pressures observed by ISCCP (Figure 5) but not the divergent treatments of partially-cloudy pixels. Robust differences between the model and observations, such as the lack of modeled clouds with $p_c > 800$ mb, can be confidently attributed to model error. 41

10 Taylor diagrams showing agreement between measures of total cloudiness (top panel) and clouds with optical thickness greater than 1.3 (bottom) between AM3's ISCCP simulator and ISCCP observations (green), AM3's MODIS simulator and MODIS observations (blue), and the two sets of observations (pink). Statistics are computed globally over the composite seasonal cycle. Observations are the reference for comparison with the model; because the reference varies among the comparisons the Taylor diagram has been non-dimensionalized. Excluding optically thin clouds from the comparison greatly reduces the bias and improves the correlation between the two sets of observations, making model evaluation more robust. 42

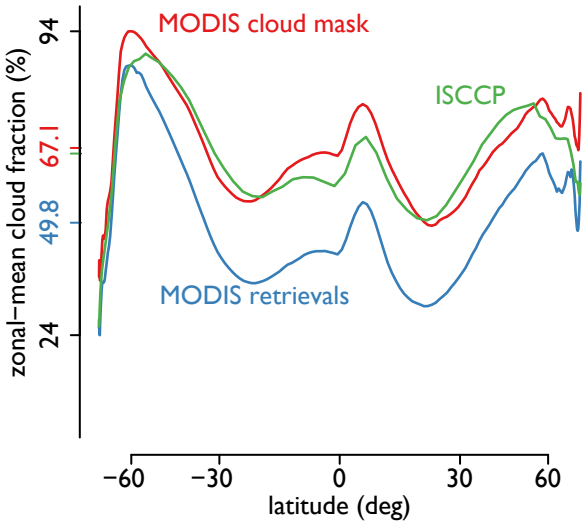

FIG. 1. Climatological zonal-mean distribution of daytime cloud fraction from ISCCP (green), the MODIS cloud mask (red), and MODIS cloud retrievals (blue, including all ice, liquid, and undetermined clouds). Tick marks on the ordinate show the corresponding global-mean cloud fraction and the range of zonally-averaged cloud fractions. ISCCP and the MODIS cloud mask produce similar distributions of cloud occurrence, while the MODIS retrievals (which exclude pixels thought to contain broken clouds) are much more conservative.

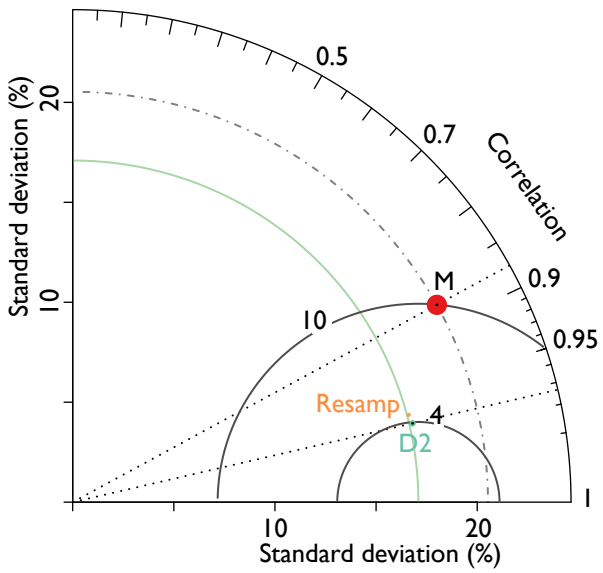

FIG. 2. Agreement in the composite annual cycle of daytime-mean climatology between daytime-only ISCCP observations and other measures of cloud fraction. Standard deviation and correlation coefficients are computed over the twelve months of the composite annual cycle, weighting spatially according to area. Climatologies created using different strategies for averaging ISCCP data over the diurnal cycle ("D2" and "Resamp") are quite similar to the daytime-only climatology, implying that differences between ISCCP and the MODIS cloud mask ("M") reflect other sampling differences, including the sampling of viewing and illumination geometry.

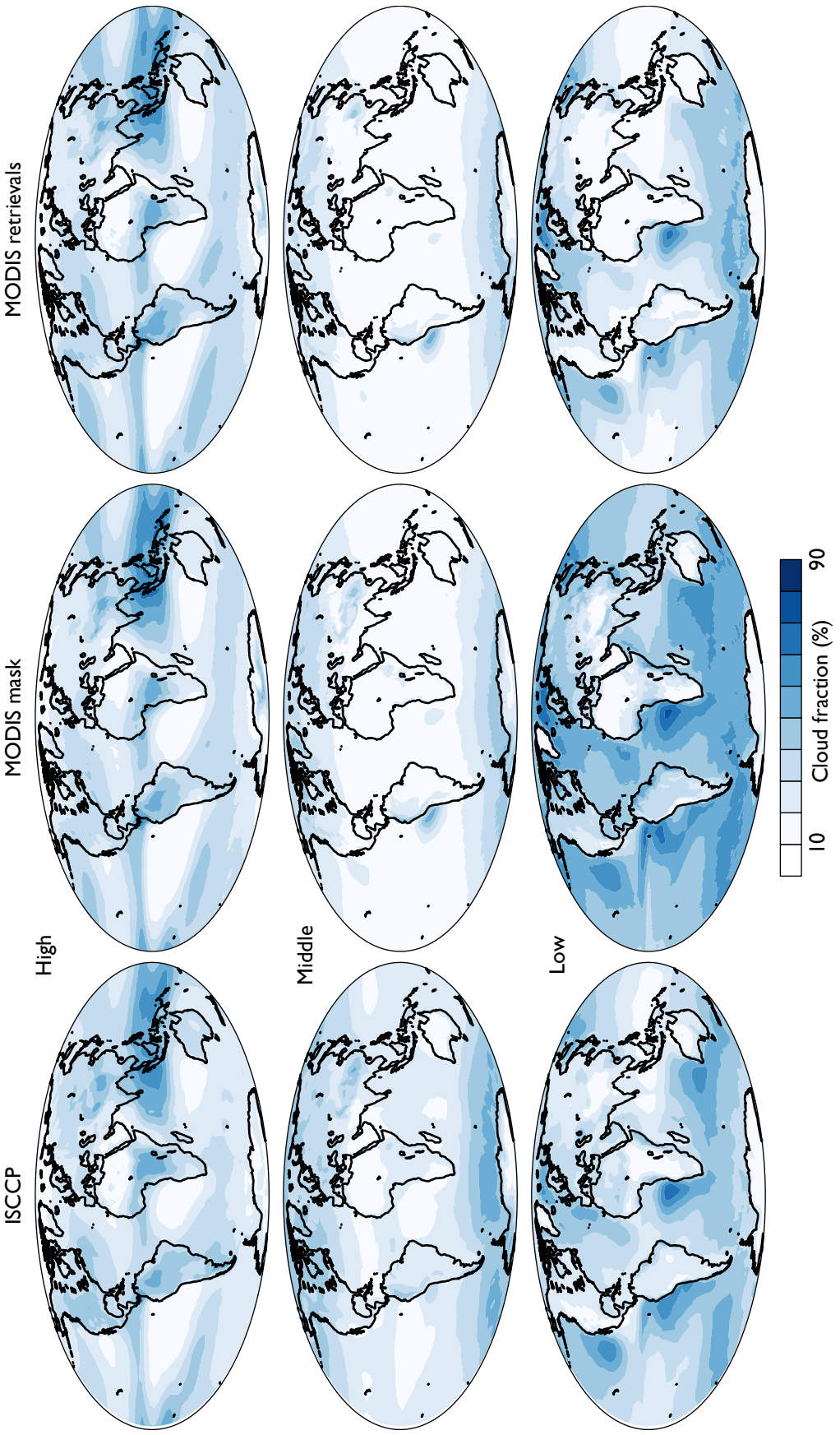

FIG. 3. Annual-mean distribution of high ($p_c < 440$ hPa, top row), middle (440 hPa $< p_c <$ 680 hPa, center), and low ($p_c > 680$ hPa, bottom) clouds obtained from ISCCP (left column), the MODIS cloud mask (center), and MODIS cloud retrievals (right). ISCCP depends on signals in the thermal infrared, and assigns clouds to mid-levels more frequently than MODIS (which uses CO_2 slicing) or other sensors. The clear-sky restoral that distinguishes the MODIS retrievals from the MODIS cloud mask is more likely to exclude low clouds than high clouds; we infer that low clouds are more spatially inhomogeneous.

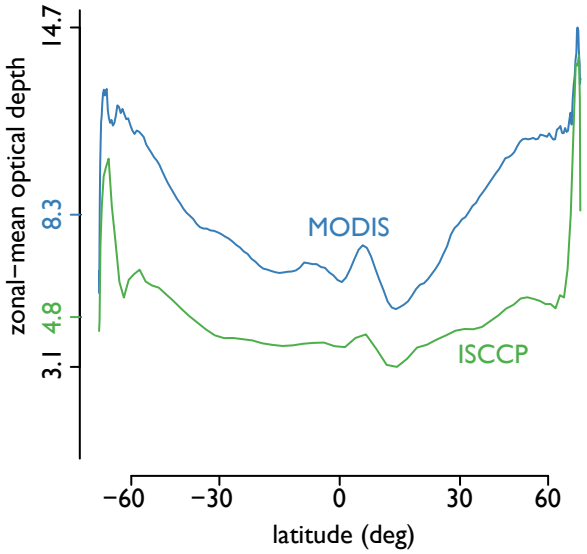

FIG. 4. Top: Climatological zonal-mean distribution of "radiatively effective" optical thickness (i.e. the optical thickness that produces the mean albedo, described in Section 2) from ISCCP (τ_r, green) and MODIS cloud (τ_l, blue). Tick marks on the ordinate show the corresponding global mean for each dataset and the overall range. Temporal averaging is uniform: monthly means are weighted by the daily faction, then combined linearly over years. MODIS optical depths are substantially higher than ISCCP because the edge detection invoked by "clear-sky restoral" preferentially removes optically thin pixels.

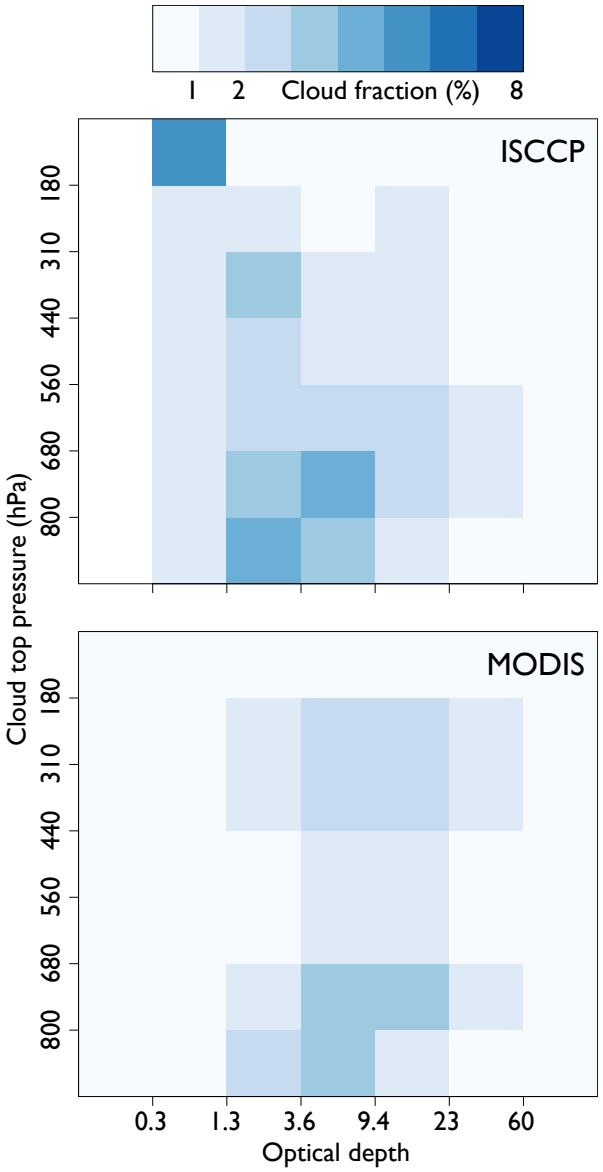

FIG. 5. Climatological distribution of global-mean cloud fraction $c_f(\tau, p_c)$ as a function of cloud top pressure p_c (vertical axis) and cloud optical thickness τ (horizontal axis) observed by ISCCP (upper panel) and MODIS (lower). ISCCP provides retrievals for all pixels identified as cloud-affected, while MODIS excludes those identified as marginal by the "clear-sky restoral" process described in the text. The broken clouds most frequently removed during clear-sky restoral represent almost all the optical thinnest clouds ($\tau < 1.3$) observed by ISCCP. The frequent observation of very high, very thin clouds by ISCCP may indicate that many of these observations are difficult to interpret self-consistently.

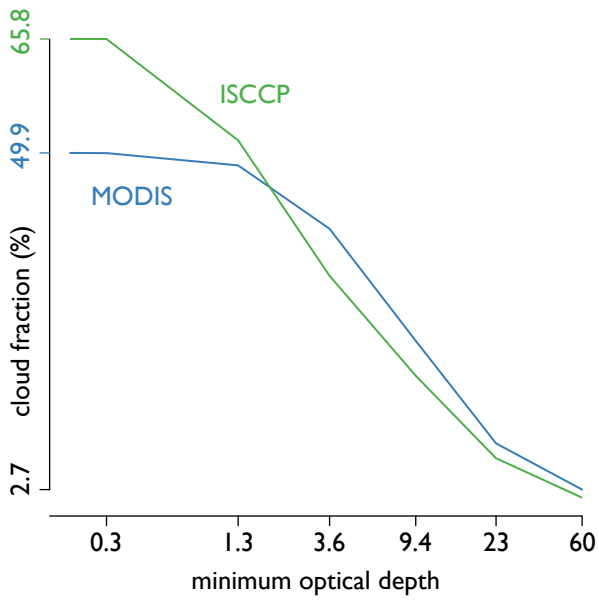

FIG. 6. Global-mean climatological values of day time cloud fraction $c_f(\tau \geq \tau_{\text{min}})$ as a function of the minimum optical depth τ_{min}. The global-mean value is computed by summing the joint histogram $c_f(x, y, t; \tau \geq \tau_{\text{min}}, p_c)$ and averaging in time and space. The large difference in cloudiness between MODIS retrievals, on the one hand, and ISCCP and the MODIS cloud mask, on the other (see Fig. 1) is due almost entirely to clear-sky restoral removing pixels with $0.3 \leq \tau \leq 1.3$; most of these are initially identified as low clouds by MODIS.

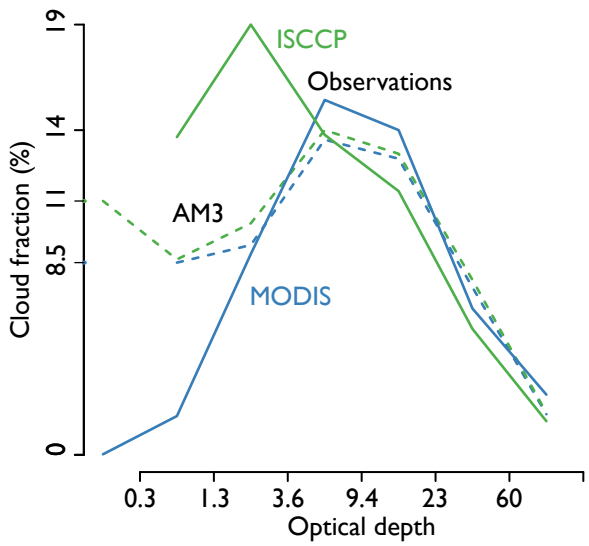

FIG. 7. Marginal histogram of global-mean cloud fraction as a function of optical thickness τ in observations (solid lines) and from instrument simulators running in the AM3 climate model (dashed). ISCCP and the ISCCP simulator are in green; MODIS and the MODIS simulator are in blue. The instrument simulators produce very similar distributions, though some MODIS simulator cloud fractions are slightly smaller at most values of τ due to failed retrievals of particle size in mixed-phase clouds. Neither simulator accounts for retrievals in scattered clouds where the observations disagree most sharply (see Section 3). Clouds with $\tau \leq 0.3$ are quite common in AM3 though they are not detectable in passive observations.

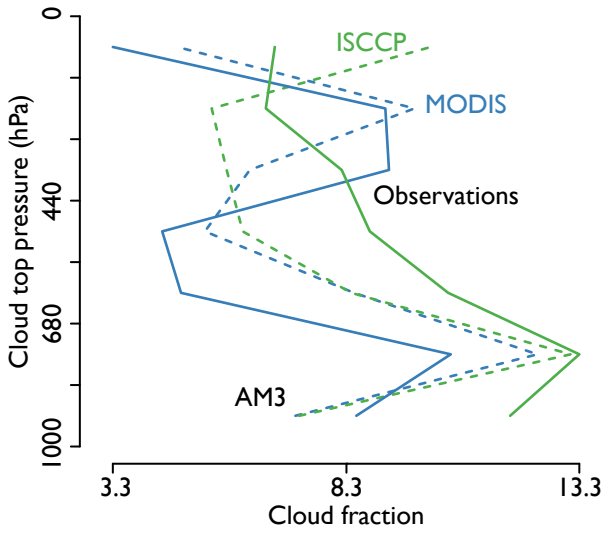

FIG. 8. Marginal histogram of global-mean cloud fraction as a function of cloud top pressure p_c in observations (solid lines) and from instrument simulators running in the AM3 climate model (dashed). ISCCP and the ISCCP simulator (excluding clouds with $\tau \leq 0.3$) are in green; MODIS and the MODIS simulator are in blue. Total observed cloud fractions are lower in MODIS than ISCCP primarily because MODIS does not aggregate marginal pixels which tend to be optically thin and low in the atmosphere. ISCCP thermal algorithms assigns far more cloudiness to the middle of the atmosphere ($p_c \leq 440 < 680$ hPa) than does MODIS's CO_2 slicing method. The MODIS and ISCCP simulators do not capture this detail, in part because the joint distribution of clouds $c_f(\tau, p_c)$ in AM3 is not the same as the distribution in nature.

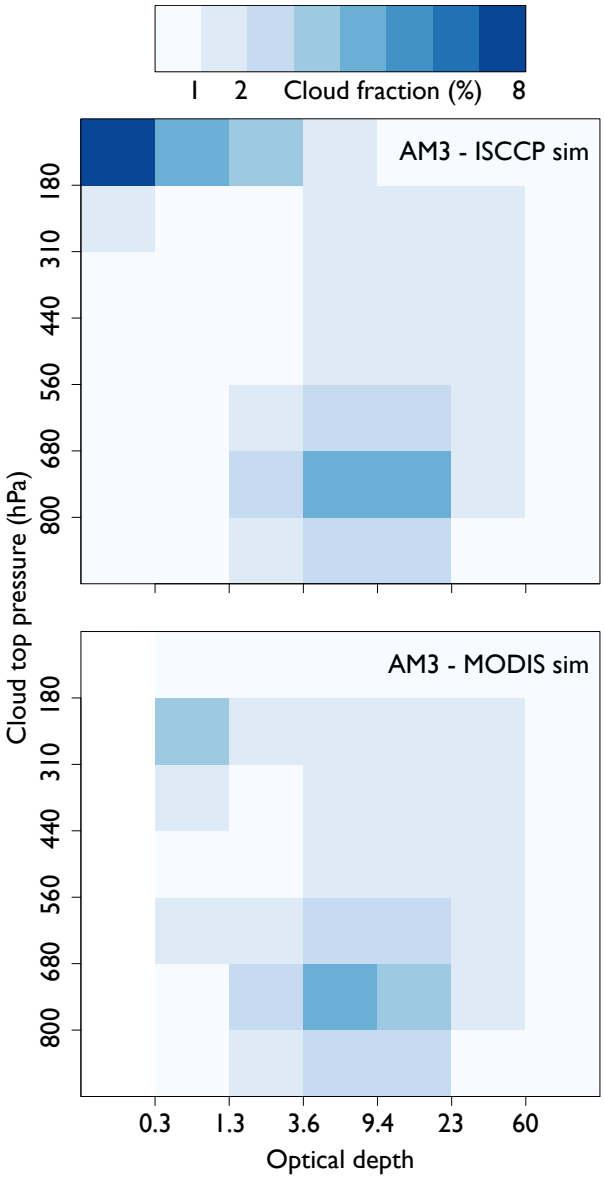

FIG. 9. Climatological distribution of global-mean cloud fraction $c_f(\tau, p_c)$ as a function of cloud top pressure p_c (vertical axis) and cloud optical thickness τ (horizontal axis) from the ISCCP (upper row) and MODIS (lower row) simulators running in the AM3 climate model. Very optically thin clouds ($\tau \leq 0.3$) cover about 11% of the earth in AM3; these are not detectable in either set of observations. The simulators reproduce the higher cloud top pressures observed by ISCCP (Figure 5) but not the divergent treatments of partially-cloudy pixels. Robust differences between the model and observations, such as the lack of modeled clouds with $p_c > 800$ mb, can be confidently attributed to model error.

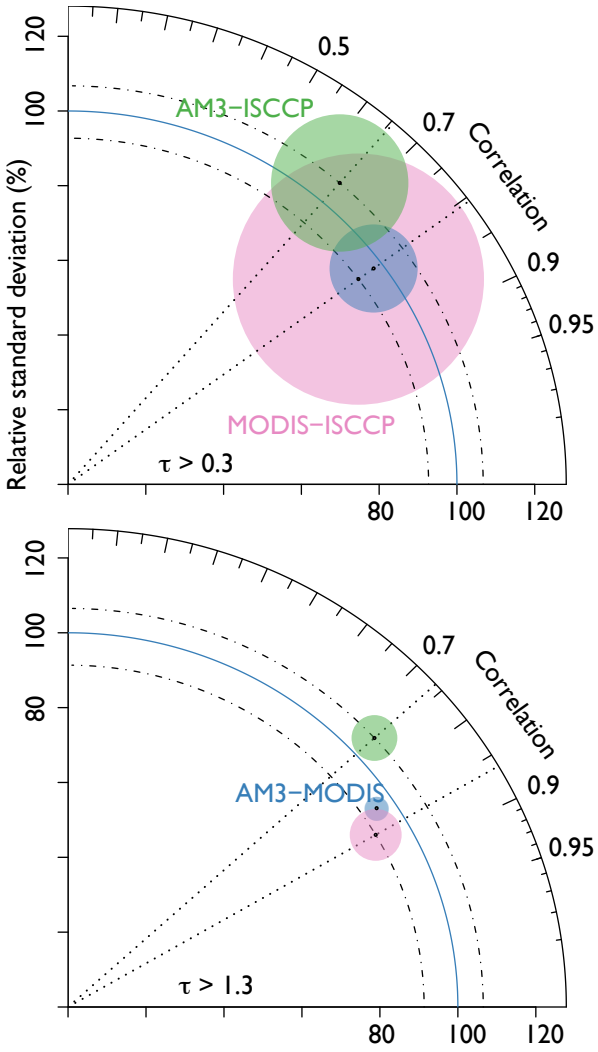

FIG. 10. Taylor diagrams showing agreement between measures of total cloudiness (top panel) and clouds with optical thickness greater than 1.3 (bottom) between AM3's ISCCP simulator and ISCCP observations (green), AM3's MODIS simulator and MODIS observations (blue), and the two sets of observations (pink). Statistics are computed globally over the composite seasonal cycle. Observations are the reference for comparison with the model; because the reference varies among the comparisons the Taylor diagram has been non-dimensionalized. Excluding optically thin clouds from the comparison greatly reduces the bias and improves the correlation between the two sets of observations, making model evaluation more robust.